Complex patterns of the HIV-1 epidemic in Kuala Lumpur, Malaysia: Evidence for expansion of circulating recombinant form CRF33_01B and detection of multiple other recombinants

Bin Wang a,1, Katherine A. Lau a,1, Lai-Yee Ong b, Meet Shah a, Megan C. Steain a, Brian Foley c, Dominic E. Dwyer d, Choo Beng Chew d, Adeeba Kamarulzaman b, Kee Peng Ng b, Nitin K. Saksena a,⁎

a Retroviral Genetics Division, Center for Virus Research, Westmead Millennium Institute, Westmead Hospital, The University of Sydney, Darcy Road, Westmead NSW 2145, Sydney, Australia
b Department of Medical Microbiology and Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, 50603, Malaysia
c HIV Sequence Database, Los Alamos, NM, USA
d Department of Virology, CIDMLS, ICPMR, Westmead Hospital, Darcy Road, Westmead NSW 2145, Australia

Received 13 February 2007; returned to author for revision 10 May 2007; accepted 15 May 2007
Available online 29 June 2007

Abstract

The HIV protease-reverse transcriptase (PR-RT) (1047 bp), gp120-env (891 bp) and gp41-env (547 bp) regions from the plasma of 115 HIV-1-infected patients in Kuala Lumpur (KL), Malaysia were sequenced. Detailed phylogenetic and bootscanning analyses were performed to determine the mosaic structure of the HIV-1 strains and their recombination breakpoint(s). Among the 50 patient samples in which all three regions could be amplified, the HIV-1 CRF01_AE subtype (46%) was predominant followed by subtypes B (10%) and B′ (6%). A total of 9/50 (18%) patients were infected with a CRF01_AE/B inter-subtype recombinant, displaying a recombinant form (RF) PR-RT, CRF01_AE gp120-env and CRF01_AE gp41-env. This RF was derived from the Thai variants of CRF01_AE and B′ subtype, with two distinct B′ subtype segments in the backbone of CRF01_AE, similar to the newly identified CRF33_01B. In addition, one sample demonstrated a close structural relationship with the new CRF33_01B in the PR-RT region but displayed B′ segment in part of the env region (RF PR-RT, CRF01_AE/B gp120-env and B′ gp41-env) indicating continuing evolution of CRF33_01B. The remaining 18% of samples were identified as unique recombinant forms (URFs).

© 2007 Elsevier Inc. All rights reserved.

Keywords: Malaysia; HIV; AIDS; Recombination; Circulating recombinant forms; Epidemiology

Introduction

As the acquired immune deficiency syndrome (AIDS) pandemic enters its third decade, the number of people living with human immunodeficiency virus (HIV) infection continues to increase. Although the HIV/AIDS epidemic was recognized in Southeast Asia later than elsewhere, local risk behaviors have allowed the epidemic to expand rapidly. Today, injecting drug use (IDU) accounts for up to 70% of HIV-1 transmission in many Asian countries, including China, Indonesia, Malaysia, Myanmar, Eastern India and Vietnam (Cock and Weiss, 2000; Saksena et al., 2005). Also, there is ample evidence that heterosexual transmission through commercial sex workers has increased over the last few years (Saksena et al., 2005).

Phylogenetic analysis shows that HIV-1 can be divided into three distinct groups: M (major), O (outlier) and N (non-M non-O or new). The M subgroup, which is associated with the global pandemic, is further subdivided into subtypes (A–D, F–H, J and K) and circulating recombinant forms (CRFs) (McCutchan, 2000; Peeters and Sharp, 2000). Previous studies indicated that B and CRF01_AE subtypes were predominant in parts of Southeast Asia. These HIV-1 strains were initially introduced independently into different risk groups, with B subtype being...
the most frequently observed among IDU and CRF01_AE being more prevalent among commercial sex workers (Ou et al., 1993). However, in recent years, there has been a dramatic shift with CRF01_AE now disseminating faster in all risk groups (Kalish et al., 1994; Weniger et al., 1994), together with the emergence and spread of new CRFs and unique recombinant forms (URFs) involving the CRF01_AE and B subtypes (Ramos et al., 2003; Tee et al., 2005; Tovanabutra et al., 2003). The emergence of CRF01_AE/B recombinants in Thailand and Malaysia is explained by CRF01_AE and B subtype co-circulation and dual infections, an ideal environment for recombination (Robertson et al., 1999). Previous studies have indicated that the pattern of the HIV-1 epidemic in Southeast Asia is complex, dynamic and with a high frequency of inter-subtype recombination (Saksena et al., 2005). Recent trends suggest a gradual replacement of ‘pure’ subtypes with recombinant HIV-1 strains in this geographical region, as evident from CRF01_AE (Tovanabutra et al., 2004). It may be that these CRFs and URFs are more ‘fit’ than their parental subtypes (Steain et al., 2004; Wang et al., 2000). However, only a limited number of these newly emerging CRF01_AE/B recombinant forms have been characterized (e.g. CRF15_01B), while the structure of other recombinant strains remains to be elucidated (Tovanabutra et al., 2003).

Surveillance and monitoring studies in Malaysia indicate that the number of new HIV/AIDS cases has risen exponentially since the first case was reported two decades ago (Huang and Hussein, 2004). In 2003, it was estimated that more than 69,000 persons were infected with HIV-1 with 76% of infections acquired through IDU and 13.4% via heterosexual transmission (UNAIDS, 2006). Similar to Thailand, CRF01_AE and B subtypes have been predominant in Malaysia, with evidence that recombination between these subtypes has occurred (Tee et al., 2005).

Here we show the evidence for the rapid expansion of a new CRF33_01B subtype in Kuala Lumpur (KL), Malaysia, based on its detection in a cluster of epidemiologically unlinked individuals. Several new URFs were also identified, present at low prevalence in various risk groups in KL.

Results

Co-circulation of multiple HIV-1 genetic subtypes and their recombinant strains based on one, two or three HIV-1 genomic regions

Fifty plasma samples were analyzed and successfully sequenced in the HIV-1 PR-RT, gp120-env and gp41-env regions. We confirmed the predominance of CRF01_AE in Malaysia, observed in 23/50 (46%) samples (Table 1). Another 5 (10%) and 3 (6%) samples were identified as subtype B and the Thai B′ subtype variant, respectively. Three samples (B_PRRT, B′gp120-env and B′gp41-env) were classified as B/B′ intra-subtype recombinant forms, with the remaining 16 (38%) samples displaying subtype discordance in the analyzed regions, suggesting possible inter-subtype recombinants.

Apart from the 50 plasma samples from which all three HIV-1 regions were sequenced, we sequenced either one or two HIV-1 genomic regions from another 65 patients from KL, for which repeated attempts failed to obtain sequences of all three genomic regions (attributed to either sequence heterogeneity imposing technical limitations, or low plasma viremia levels in patients on highly active antiretroviral therapy). For 22 samples where only two HIV-1 genomic regions were successfully sequenced, we found evidence for the co-circulation of a variety of HIV recombinants and pure subtypes. The majority of these samples were CRF01_AE and B subtypes, based on either PR-RT and gp120-env, or PR-RT and gp41-env sequences. We identified a number of novel unique recombinant forms, consisting of CRF02_AG, CRF07_BC, CRF08_BC, CRF14_BG and CRF15_01B (data not shown—refer to appended Supplementary data 1).

Additionally, there were 43 samples where only a single HIV-1 genomic region was sequenced in either PR-RT or one of the two env regions (gp120-env or gp41-env). Based on these data, there was evidence for the presence of a number of URFs in low prevalence, such as CRF03_AB, CRF08_BC, CRF14_BG and CRF15_01B. There was evidence of HIV-1 subtypes C and D, which appear to be uncommon in this country (data not shown—refer to appended Supplementary data 1).

Identification of HIV-1 subtype CRF33_01B and its molecular and epidemiological characteristics

The regions sequenced in PR-RT, gp120-env and gp41-env were 2252–3299, 6867–7757 and 7758–8308 respectively. Ten (20%) of 50 samples displayed a similar recombination pattern within the PR-RT gene region and were similar to CRF01_AE (Fig. 1). However, using the bootscan approach, three recombination breakpoints were seen, revealing two distinct B′ subtype segments within the backbone of CRF01_AE (Fig. 2A). Phylogenetic neighbor-joining trees of partial segments delimited by consensus breakpoints (i.e. the range of nucleotide positions which coincides in all 10 samples) (Figs. 3A–D) confirm the subtype assignments of segments obtained by boot-
scanning. The PR-RT region can therefore be distributed into four different segments; the 1st and 3rd segments clustered with B′ subtype, while the 2nd and 4th segments clustered with CRF01_AE. Our analysis using bootscanning and phylogenetic reconstructions based on a sliding window identified considerable heterogeneity in breakpoints of these samples, depicted in Table 2. Due to high genetic similarities in the genomic regions among subtypes B′ and CRF01_AE, the exact nucleotide position at which the breakpoints occurred could not be determined. Three relative breakpoint regions were identified at positions 2393–2462 nt, 2529–2549 nt and 2847–2848 nt (nucleotide numbering relative to HXB2; http://www.hiv.lanl.gov/content/hiv-db/LOCATE/locate.html). Simplot analysis also further confirmed the high similarity of these ten samples with the newly identified CRF33_01B (data not shown—refer to appended Supplementary data 2).

Further analysis of the gp120-env and gp41-env regions showed that although all 10 samples displayed similar genome structure in the PR-RT gene region with CRF33_01B, only 9 exhibited a consistent genome structure within the gp120-env and gp41-env regions. Sample 06MYKLD46 showed a distinct pattern from the other 9 samples, with a distinct breakpoint detected approximately at 7550 nt (Fig. 2B) in the continuous region of gp120-env and gp41-env (from 6867 to 8304 nt). Despite the presence of CRF01_AE sequence at the 5′ sub-region (683 bp) prior to the breakpoint, a 754 bp B′ subtype segment was inserted after the breakpoint (Table 2).

The breakpoint distributions and the resulting subtype variability across the three HIV-1 genome regions in the described recombinant forms were compared with the Thai CRF01_AE (AY358068), B (AY945711) and CRF01_AE/B (97TH.NP1623, 99TH.MU2079, 97THR01, 96THR02, 97THR03 and 95THR04) and 05MYKL007.1 (CRF33_01B) (Fig. 4). Sample 06MYKLD46 and the 9 CRF33_01B samples showed two distinct B′ subtype segments in a backbone of CRF01_AE in the PR-RT genomic region. Among all the CRF01_AE/B Thai variants, only the 96THR02 and 97THR03 strains showed a similar structural relationship within the analyzed PR-RT gene region to the described 10 samples. Both of these strains have two short segments of B′ subtype-like sequence close to the 5′ sub-region, but exhibit different recombinant breakpoints. As for gp120-env and gp41-env regions, only sample 06MYKLD46 exhibited a downstream subtype B′ insertion from 7551 to 8304 nt. In contrast, none of the recombinant strains displayed B′ subtype in these regions, except 97TH.NP1623 and 99TH.MU2079, which had B′ subtype in the gp41-env only, and entire gp120-env and gp41-env gene regions, respectively.
All 10 samples were examined to identify any possible epidemiological association. We found that this novel CRF33_01B and sample 06MYKLD46 were relatively common in IDU (5/10), followed by heterosexuals (3/10), hetero-IDU (1/10) and bisexuals (1/10) (Table 3).

Evidence for continuing evolution and recombination event of CRF33_01B

The genome topology of this newly identified CRF (RFPR-RT, CRF01_AE gp120-env and CRF01_AE gp41-env) in 9/50 patients illustrated an insertion of two short segments of B’ subtype in the PR-RT region of the backbone of CRF01_AE. Concurrently, we found an extra subtype B’ fragment at the 3’ end of gp120-env region (from position 7551 to 7757 nt), which then continues into gp41-env, in sample 06MYKLD46, which sets it apart from CRF33_01B. This particular sample (RFPR-RT, CRF01_AE/B gp120-env and B’ gp41-env) demonstrated a close structural relationship with the CRF33_01B recombinant.

Possible recombination hotspots in CRF01_AE/B’

There were three other samples (06MYKL007, 06MYKL017, 06MYKL145), which showed CRF01_AE/B’ recombinant pattern in the PR-RT gene region. (A) Sample 06MYKL044 was a representative sample of the group of 10 identical URF in PR-RT gene region. Three recombinant breakpoint regions were identified at 2393–2462 nt, 2529–2549 nt and 2847–2848 nt (HXB2 numbering). (B) Sample 06MYKLD46 of entire gp120-env and gp41-env gene regions. A recombinant breakpoint at approximately 7550 nt was identified. (C) Sample 06MYKL007 of CRF01_AE/B recombinant pattern in the PR-RT gene region. (D) Sample 06MYKL017 with an approximate breakpoint at 3172 nt. (E) Sample 06MYKL145 with an approximate breakpoint at 2982 nt.

Fig. 2. Bootscanning plots of HIV-1 gene sequences of HIV-1 isolate from Kuala Lumpur, Malaysia. The subtype B’ reference strain (AY713408) and Thai CRF01_AE reference strain (AY358068) were used as putative parental strains, with subtype C (AB023804) used as an out group. The bootstrap values were plotted for a window of 200 bp moving in increment of 20 bp along with the alignment. Y-axis represents the % of permuted trees. (A) Sample 06MYKL044 was a representative sample of the group of 10 identical URF in PR-RT gene region. Three recombinant breakpoint regions were identified at 2393–2462 nt, 2529–2549 nt and 2847–2848 nt (HXB2 numbering). (B) Sample 06MYKLD46 of entire gp120-env and gp41-env gene regions. A recombinant breakpoint at approximately 7550 nt was identified. (C) Sample 06MYKL007 of CRF01_AE/B recombinant pattern in the PR-RT gene region. (D) Sample 06MYKL017 with an approximate breakpoint at 3172 nt. (E) Sample 06MYKL145 with an approximate breakpoint at 2982 nt.
Fig. 3. Neighbor-joining phylogenetic trees of partial segments of recombinant CRF01_AE-B viruses from Kuala Lumpur, Malaysia. Nucleotide positions (HXB2 numeration) that delimit the analyzed fragments and the segment number according to Table 2 are shown at the bottom of each tree: (A) Segment 1. (B) Segment 2. (C) Segment 3. (D) Segment 4. Trees are rooted with subtype C (AB023804). Subtype B reference sequences are AY173951, AY945711, AY713408 and AY945710, and CRF01_AE reference sequences are AF259955, AY358067, AB032741 and AB032740. Each reference sequence is labeled with their accession number followed by their HIV-1 subtypes. Bootstrap values ≥70% (based on 100 replicates) are shown.
patterns within the PR-RT gene region (Figs. 2C–E). They all contained CRF01_AE in both gp120-env and gp41-env gene regions. However, for the sample 06MYKL007, one breakpoint was detected at position 2393–2462 nt of the PR-RT region, similar to the first breakpoint of the newly described CRF, with a short segment of B' subtype (140 bp) at the 5' sub-region of CRF01_AE backbone. For sample 06MYKL017, there was a considerably long segment of B' subtype (920 bp) prior to the breakpoint at approximately 3172 nt, while the 127 bp 3' segment of the PR-RT region was predicted as CRF01_AE. Finally, the PR-RT gene region for sample 06MYKL145 was estimated as CRF01_AE until position 2982 nt followed by B' subtype from 2983 to 3299 nt. This higher prevalence of genetic breakpoints in the HIV-1 PR-RT region may indicate a possible preference region for recombination to occur.

Emergence of new URFs

We also found evidence of HIV-1-infected patients harboring new URFs in Malaysia. Samples 06MYKLU6, 06MYKL005 and 06MYKL077 were deduced to have subtype discordance encompassing the three distinct HIV-1 genomic regions included in this study. Sample 06MYKLU6 was a CRF01_AEPR-RT, CRF01_AEgp120-env and B'gp41-env recombinant, while sample 06MYKL005 was predicted to have B'PR-RT, CRF01_AE/gp120-env and B'gp41-env (Fig. 4). In 06MYKL005, the single recombinant breakpoint at gp120-env was at approximately 7550 nt, similar to that observed in 06MYKLD46. The other URF was observed in sample 06MYKL077, whereby gp41-env showed CRF09_cpx sequence while both PR-RT and gp120-env regions consisted of a sequence of unknown HIV-1 genetic subtype closely related to CRF02_AG, CRF09_cpx and A subtype (data not shown, refer to appended Supplementary data 3).

Overall, the URFs (06MYKLU6, 06MYKL005, 06MYKL077, 06MYKLU07, 06MYKL017 and 06MYKL145), along with B'/B' intra-subtype recombinants, were prevalent in 12% of heterosexuals. Their low prevalence in other risk groups such as bisexuals and men who have sex with men suggests their gradual dissemination and expansion in all other HIV-risk groups.

Discussion

A unique trait of HIV is its ability to generate multiple genetic variants that may be more fit and virulent (Saksena et al., 2005; Stein et al., 2004). Recombination is one strategy HIV employs to generate genetically diverse strains; there is only partial biological proof to attest the belief (Wang et al., 2000) that these recombinant forms have altered cytopathic and/or pathogenic potential. Those recombinant HIV-1 strains that have disseminated widely can be defined as CRFs (Carr et al., 1998). They originated from Central Africa (Carr et al., 1996; Gao et al., 1996; Murphy et al., 1993) and has spread throughout Southeast Asia (Kalish et al., 1994) and has spread throughout Southeast Asia (Kalish et al., 1994) and has spread throughout Southeast Asia (Kalish et al., 1994) and has spread throughout Southeast Asia (Kalish et al., 1994). Many regions in Southeast Asia have experienced the dispersal of multiple HIV-1 subtypes and the emergence of a variety of inter-subtype recombinants, suggesting that the HIV-1 epidemic in Asia is in a dynamic state of evolution.

Table 2
Mosaic structure of recombinant form of HIV-1 patient samples in Kuala Lumpur, Malaysia

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>HXB2 numbering (nt)</th>
<th>Segment breakpoint</th>
<th>Gene region</th>
<th>Sequence range (nt)</th>
<th>Sequence length</th>
<th>Subtype</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2252–3299</td>
<td>1st segment</td>
<td>PR</td>
<td>2253–2392</td>
<td>140</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1st break</td>
<td>PR</td>
<td>2393–2462</td>
<td>70</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2nd segment</td>
<td>PR-RT</td>
<td>2463–2528</td>
<td>66</td>
<td>CRF01_AE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2nd break</td>
<td>RT</td>
<td>2529–2549</td>
<td>21</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3rd segment</td>
<td>RT</td>
<td>2550–2848</td>
<td>297</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3rd break</td>
<td>RT</td>
<td>2847–2848</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4th segment</td>
<td>RT</td>
<td>2849–3299</td>
<td>451</td>
<td>CRF01_AE</td>
<td></td>
</tr>
<tr>
<td>6867–7775</td>
<td>1st segment</td>
<td>gp120-env</td>
<td>6867–7757</td>
<td>891</td>
<td>CRF01_AE</td>
<td></td>
</tr>
<tr>
<td>7758–8304</td>
<td>1st segment</td>
<td>gp41-env</td>
<td>7758–8304</td>
<td>547</td>
<td>CRF01_AE</td>
<td></td>
</tr>
<tr>
<td>06MYKLD46</td>
<td>2252–3299</td>
<td>1st segment</td>
<td>PR</td>
<td>2253–2392</td>
<td>140</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>1st break</td>
<td>PR</td>
<td>2393–2462</td>
<td>70</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2nd segment</td>
<td>PR-RT</td>
<td>2463–2528</td>
<td>66</td>
<td>CRF01_AE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2nd break</td>
<td>RT</td>
<td>2529–2549</td>
<td>21</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3rd segment</td>
<td>RT</td>
<td>2550–2848</td>
<td>297</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3rd break</td>
<td>RT</td>
<td>2847–2848</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4th segment</td>
<td>RT</td>
<td>2850–3299</td>
<td>451</td>
<td>CRF01_AE</td>
<td></td>
</tr>
<tr>
<td>6867–7775</td>
<td>1st segment</td>
<td>gp120-env</td>
<td>6867–7549</td>
<td>683</td>
<td>CRF01_AE</td>
<td></td>
</tr>
<tr>
<td>7758–8304</td>
<td>1st break</td>
<td>gp120-env</td>
<td>7550</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2nd segment</td>
<td>gp120-env</td>
<td>7551–7757</td>
<td>207</td>
<td>B</td>
<td></td>
</tr>
</tbody>
</table>
| Note. PR, protease; RT, reverse transcriptase.
Here, we have identified and characterized complex HIV-1 genomes from Kuala Lumpur, Malaysia. Although CRF01_AE is still the major circulating recombinant form, analysis of samples collected in 2005–2006 shows that CRF33_01B, a CRF01_AE/B inter-subtype recombinant form, is emerging in Malaysia, consistent with recent findings showing the analysis of HIV-1 strains that were collected in 2002 (Tee et al., 2006). The prevalence of CRF33_01B was found in 18% of the HIV-1-infected population in KL, Malaysia, suggesting its active circulation and expansion in various HIV-I infection risk groups. CRF33_01B seems to have broken the transmission barrier from IDU to heterosexual, bisexual and hetero-IDU groups, highlighting the complexity of transmission patterns and possible dissemination routes of this novel form into the general population. It also suggests that this new CRF is certainly competing actively with other circulating HIV-1 subtypes and CRFs in the region. Interestingly, our evidence of continuing evolution and recombination in CRF33_01B with the emergence of 2nd generation recombinants further highlights the increasing complexity of the HIV epidemic in this area. Similar evidence for the emergence of 2nd and 3rd generation recombinants has been observed in China (Saksena et al., 2005; Yang et al., 2003).

From comparison of the present and previous studies of HIV-1 epidemic in China and Southeast Asia, it is apparent that the rapid dissemination and establishment of certain recombinant viral strains may be gradually phasing-out ‘pure’ HIV-1 subtypes.
The drivers of both the Chinese and Southeast Asian epidemics are predominantly HIV-1 inter-subtype recombinants (Yang et al., 2002, 2003). For example, in Yunnan province of China, B’ subtype has actively replaced the B subtype of American lineage, increasing from 20% in 1990 to 90% in 1996 (Yang et al., 2003). Following this, HIV-1 C subtype strains were also identified in the early 1990s among IDUs (Luo et al., 1995). As a result of the co-circulation of the B and C subtypes, two CRFs arose, CRF07_BC and CRF08_BC, most likely in this province among IDUs (Piyasirisilp et al., 2000; Su et al., 2000). Consistent with the emergence of various recombinants in Southeast Asia, we have identified a variety of new CRFs at low prevalence. These have gradually emerged in high risk groups, possibly as a consequence of dual and superinfections. Dual infection or superinfection may result from the simultaneous passage of multiple HIV strains either during a single transmission event or from the sequential passage of viruses during multiple transmission events (Steain et al., 2004). Despite limited number of these cases identified, the existence of multiple subtypes in a given geographical region strengthens the likelihood of both scenarios.

To date, recombinant viruses have been estimated to account for about 20% of all HIV infection in some countries (Neilson et al., 1999), but the biological consequences of recombination and superinfection have not been fully elucidated. There is limited information on the importance of CRFs phenotypes in cell tropism, drug resistance, replication kinetics, transmission rate and disease progression.

We emphasize that the increase in the prevalence of newly emerging CRF01_AE/B indicates continued evolution, which may have resulted in altered virologic characteristics and possible transmission advantages. A snapshot of this is apparent from our study on the description and expansion of the new CRF33_01B recombinant. Given the high plasticity of the HIV genome and the burgeoning nature of economies in China, India, Thailand, Malaysia and Russia, such trends may become more common. This may further complicate the timely diagnosis of molecularly altered forms of HIV, which may be capable of causing epidemiologic shifts. Continued molecular characterization and epidemiological monitoring of these recombinant forms is needed. Our study has documented both the expansion of new CRF33_01B and the emergence of several new URFs, which may be possible contributors in shaping the nature of HIV epidemic in Malaysia.

Materials and methods

Clinical samples

Plasma samples from 115 HIV-1 seropositive patients were collected from the University Malaya Medical Center (UMMC) HIV Clinic during 2005 to 2006 and stored at −80 °C. The major ethnic groups in this study were Chinese (44%), Malay (30%), Indian (16%) and others (10%). This study was approved by the UMMC Medical Ethics Committee and informed consent was obtained from all participants or their families prior to sample collection.

Phylogenetic analysis

Sequences were aligned with various HIV-1 reference subtypes and CRFs obtained from the Los Alamos HIV Database (http://hiv-web.lanl.gov/), using CLUSTAL-W (Thompson et al., 1994) from the GCG package followed by minor manual

PR-RT, gp120-env and gp41-env nucleotide sequence analyses

HIV-1 RNA was extracted using the column purification method (QIAamp Viral Mini Kit, Qiagen, Hilden, Germany) and used in a nested PCR. The first-round PCR was performed according to the QIAGEN OneStep RT-PCR protocol (QIAGEN OneStep RT-PCR Kit, Qiagen, Hilden, Germany), with cycling conditions of reverse transcription at 50 °C for 30 min, followed by initial PCR activation at 95 °C for 15 min and 35 cycles of denaturation at 94 °C for 30 s, annealing at 55 °C for 30 s, extension at 72 °C for 2 min and a final extension at 72 °C for 8 min. Subsequently, a 10 μl aliquot of the first-round PCR product was subjected to a second-round PCR. For the PR-RT gene, primers PR-RT-A (nt 2025–3299) and PR-RT-4 (nt 3325–3304). For the gp120-env gene, an approximate 891 bp or 671 bp region was examined. External PCR primers used were ED5 (nt 6557–6582) and E1551R (nt 7849–7833) and internal PCR primer pairs were E615 (nt 6855–6876) and VSB (nt 7813–7789), or ES7 (nt 7002–7021) and ES8 (nt 7668–7648). Primers GP41–5 (nt 7715–7736) and GP41–4 (nt 8815–8797) were used to amplify the gp41-env gene for the first round of PCR followed by GP41–5 (nt 7715–7736) and GPAI4 (nt 8304–8285) to amplify 547 bp in the second round. Primer sequences were as previously published (Shah et al., 2006; Smit et al., 2004). The PCR amplification conditions were as follows: initial denaturation at 95 °C for 5 min, followed by 35 cycles of denaturation at 94 °C for 30 s, annealing at 55 °C for 30 s, extension at 72 °C for 2 min and a final extension at 72 °C for 8 min. PCR products were separated on 1.5% (w/v) agarose gels containing ethidium bromide (0.5 μl/ml) and the bands were visualized under UV light.

Cloning of PCR products and sequencing

PCR amplicons from the PR-RT, gp120-env and gp41-env regions were ligated into the PGEM-T Easy Vector System II (Promega, Madison, USA) and transformed into competent Escherichia coli JM109 cells, according to the manufacturer’s protocol. In case of multiple bands, band of the predicted size was excised from the gel and extracted before ligation according to the manufacturer’s instructions (QIAquick Gel Extraction Kit, Qiagen, Hilden, Germany). Screening for the gene insert was performed by quick lysis at 95 °C of the E. coli cells for 5 min followed by PCR using 5 μl of the cell lysate (Potter et al., 2006). All PCR products were purified and sequenced using the BigDye Terminator Cycle Sequencing Ready Reaction kit (ABI Prism, v 3.0) on an ABI 377 automated sequencer.
adjustments considering protein sequences. Phylogenetic trees were then constructed using the neighbor-joining method (Felsenstein, 1989). The reliability of the nodes was confirmed by bootstrap analysis with 100 replicates using the PHYLIP package (Felsenstein, 1989), and bootstrap values of ≥70% were considered to be significant for assignment of parenthood. Pairwise nucleotide evolutionary distances were calculated using the Kimura two-parameter model in PHYLIP (Kimura, 1980) with a transition/transversion ratio of 2.0.

Bioinformatic analysis of recombinants

Gaps (insertions/deletions) were stripped prior to bootscan analysis. Simplot, version 3.5 (Lole et al., 1999), was used to identify recombination breakpoints in samples with undetermined subtypes or CRF. This analysis was done by comparing each sequence against the background of reference sequences of different subtypes or CRFs selected from the Los Alamos HIV Database. Sub-region confirmatory tree analyses were carried out to confirm the subtype within each fragment and a bootstrap of ≥70% was considered to be definitive.

Acknowledgments

This work was funded by the SAGA grant 66-02-03-0043, Kuala Lumpur, Malaysia. We thank the staff of the HIV Clinic at the University Malaya Medical Centre, Kuala Lumpur, Malaysia and all the participants who volunteered for these studies. KAL was supported by the IPRS Ph.D award of The University of Sydney.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.virol.2007.05.033.

References


length human immunodeficiency virus type 1 genome of a prevalent intersubtype (C/B') recombinant strain in China. J. Virol. 74 (23), 11367–11376.


